GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        January  3, 2002, 21:32:44 ; Search time 38.81 Seconds
                                             (without alignments)
                                             1761.650 Million cell updates/sec

Title:         US-09-497-822A-19
Perfect score: 4912
Sequence:      1 MEVQLGLGRVYPRPPSKTYR.........SVQVPKILSGKVKPIYFHTQ 923

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:      522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters:      522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     A_Geneseq_1101:*
            1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT:*
            2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT:*
            3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT:*
            4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT:*
            5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT:*
            6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT:*
            7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT:*
            8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT:*
            9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT:*
           10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT:*
           11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT:*
           12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT:*
           13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT:*
           14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT:*
           15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT:*
           16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT:*
           17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT:*
           18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT:*
           19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT:*
           20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT:*
           21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT:*
           22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4880 | 99.3 | 919 | 18 | AAW14783 | Androgen receptor. |
| 2 | 4880 | 99.3 | 919 | 21 | AAY78914 | Human androgen rec |
| 3 | 4872 | 99.2 | 919 | 10 | AAP93109 | Human androgen rec |
| 4 | 4838.5 | 98.5 | 918 | 20 | AAY33491 | Human androgen rec |
| 5 | 4827 | 98.3 | 918 | 12 | AARI2223 | Human androgen rec |
| 6 | 4814 | 98.0 | 919 | 10 | AAP90996 | Human androgen rec |
| 7 | 4201.5 | 85.5 | 902 | 10 | AAP91006 | Rat androgen recep |
| 8 | 4200.5 | 85.5 | 902 | 12 | AARI2224 | Rat androgen recep |
| 9 | 4187.5 | 85.3 | 902 | 12 | AARI2224 | Rat androgen recep |
| 10 | 2429 | 49.5 | 452 | 20 | AAY21627 | Ligand binding dom |
| 11 | 1459 | 29.7 | 630 | 12 | AARI2230 | TrpE/androgen rece |

| 12 | 1349.5 | 27.5 | 260 | 22 | AAB97073 | Rat androgen recep |
| 13 | 1337.5 | 27.2 | 576 | 12 | AARI2229 | TrpE/androgen rece |
| 14 | 1261 | 25.7 | 933 | 20 | AAY21621 | Ligand binding dom |
| 15 | 1256 | 25.6 | 933 | 21 | AAY97297 | Human progesterone |
| 16 | 1071.5 | 21.8 | 984 | 20 | AAP80927 | Ligand binding dom |
| 17 | 1069.5 | 21.8 | 984 | 9 | AAP80927 | Sequence of the hu |
| 18 | 1066 | 21.7 | 1070 | 18 | AAW17789 | Green fluorescent |
| 19 | 1042 | 21.2 | 777 | 9 | AAP80919 | Sequence of the pr |
| 20 | 1042 | 21.2 | 777 | 20 | AAY21623 | Ligand binding dom |
| 21 | 1039 | 21.2 | 777 | 18 | AAW44700 | Mutant nuclear glu |
| 22 | 838 | 17.1 | 158 | 21 | AAY78913 | Androgen independe |
| 23 | 795 | 16.2 | 154 | 20 | AAY33504 | Human unliganded a |
| 24 | 765 | 15.6 | 251 | 21 | AAB26786 | Human progesterone |
| 25 | 765 | 15.6 | 255 | 22 | AAB97074 | Rat progesterone r |
| 26 | 675.5 | 13.8 | 667 | 18 | AAW44701 | Chimeric Cre-LDB-G |
| 27 | 671.5 | 13.7 | 613 | 22 | AAB47153 | iCre-FRT-GBD*.  Sy |
| 28 | 657.5 | 13.4 | 251 | 21 | AAB26787 | Glucocorticoid rec |
| 29 | 644.5 | 13.1 | 534 | 17 | AAW03197 | Glucocorticoid rec |
| 30 | 644 | 13.1 | 284 | 21 | AAB19249 | A rat glucocortico |
| 31 | 636 | 12.9 | 284 | 19 | AAW45514 | Glucocorticoid rec |
| 32 | 636 | 12.9 | 284 | 21 | AAW39140 | Rat glucocorticoid |
| 33 | 636 | 12.9 | 284 | 21 | AAY67559 | Rat glucocorticoid |
| 34 | 636 | 12.9 | 284 | 21 | AAY78888 | Glucocorticoid rec |
| 35 | 624.5 | 12.7 | 240 | 21 | AAB26781 | Glucocorticoid rec |
| 36 | 624.5 | 12.7 | 240 | 21 | AAB26785 | Glucocorticoid rec |
| 37 | 621.5 | 12.7 | 239 | 21 | AAB26788 | Glucocorticoid rec |
| 38 | 600 | 12.2 | 442 | 12 | AARI2231 | TrpE/AR androgen-b |
| 39 | 481 | 9.8 | 595 | 20 | AAY21626 | Ligand binding dom |
| 40 | 481 | 9.8 | 595 | 22 | AAG84505 | Human oestrogen re |
| 41 | 481 | 9.8 | 595 | 22 | AAG84506 | Human oestrogen re |
| 42 | 481 | 9.8 | 595 | 22 | AAG84507 | Human oestrogen re |
| 43 | 481 | 9.8 | 595 | 22 | AAG84511 | Human oestrogen re |
| 44 | 481 | 9.8 | 595 | 22 | AAG84513 | Human oestrogen re |
| 45 | 480 | 9.8 | 595 | 22 | AAG84509 | Human oestrogen re |

ALIGNMENTS

RESULT 1
AAW14783
ID   AAW14783 standard; Protein; 919 AA.
XX
AC   AAW14783;
XX
DT   22-JUN-1997  (first entry)
DE
XX
DE   Androgen receptor.
XX
KW   Androgen receptor; acidic fibroblast growth factor; aFGF;
KW   antisense; benign prostatic hyperplasia; prostate cancer; therapy.
XX
OS   Homo sapiens.
XX
PN   WO9711170-A1.
XX
PD   27-MAR-1997.
XX
PF   20-SEP-1996;  96WO-US15081.
XX
PR   20-SEP-1995;  95US-0004018.
XX
PA   (WORC-) WORCESTER FOUND BIOMEDICAL RES.
XX
PI   Hamecnik PA;
XX
DR   WPI; 1997-202879/18.
DR   N-PSDB; AAT63407.
XX
PT   Oligonucleotide(s) antisense to human androgen receptor and acidic
PT   FGF genes - used to inhibit gene expression, for the treatment of
PT   benign prostatic hyperplasia
XX

```
PS    Disclosure; Page 22-28; 51pp; English.
XX
CC    Human androgen receptor (AAW14783) binds testosterone and, acting
CC    at the transcriptional level, regulates the growth of normal
CC    prostatic cells. Antisense oligonucleotides (see also AAT63200,
CC    AAT63404-05) based on an androgen receptor cDNA clone (see also
CC    AAT63407) can be used to prevent or survival of prostatic cells for
CC    thereby inhibiting the growth or survival of prostatic cells for
CC    the treatment of benign prostatic hyperplasia and prostate cancer.
XX
SQ    Sequence    919 AA;

Query Match             99.3%;    Score 4880;  DB 18;  Length 919;
Best Local Similarity   99.6%;    Pred. No. 0;
Matches   919;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy      1 MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGPRHPEAASAAPPGASLLLLQQQ 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mevqlglgrvyprppsktyrgafqnlfqsvreviqnpgprhpeaasaappgaslll--- 57

Qy     61 QQQQQQQQQQQQQQQQQQQQETSPRQQQQQQGEDGSPQAHRRGPTGYLVLDEEQQPSQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58 -qqqqqqqqqqqqqqqqqqqetsprqqqqqqgedgspqahrrgptgylvldeeqqpsq 116

Qy    121 PQSALECHPERGCVPEPGAAVAASKLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSCCS 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    117 pqsalechpergcvpepgaavaaskqlpqqlpappdeddsaapstlsllgptfpglsscs 176

Qy    181 ADLKDILSEASTMQLLQQQQQEAVSEGSSSGRAREASGAPTSSKDNYLGGTSTISDNAKE 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    177 adlkdilseastmqllqqqqqeavsegsssgrareasgaptsskdnylggtstisdnake 236

Qy    241 LCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLLDDS 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    237 lckavsvsmglgvealehlspgeqlrgdcmyapllgvppavrptpcaplaeckgslldds 296

Qy    301 AGKSTEDTAEYSPFKGGYTKGLEGESLGCSGSAAAGSSGTLELPSTLSLYKSGALDEAAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    297 agkstedtaeyspfkggytkglegeslgcsgsaaagssgtlelpstlslyksgaldeaaa 356

Qy    361 YQSRDYYNFPLALAGPPPPPPPPHPHARIKLENPLDYGSAWAAAAAQCRYGDLASLHGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    357 yqsrdyynfplalagpppppppphphariklenpldygsawaaaaaqcrygdlaslhgag 416

Qy    421 AAGPGSGSPSAAASSSWHTLFTAEEGQLYGPCGGGGGGGGGGGGGGGGGGGGGGGGEAGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    417 aagpgsgspsaaassswhtlftaeegqlygpcgggggggggggggggggggggggeaga 476

Qy    481 VAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTCVKSEMGPWMDSYSGPYGD 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 vapygytrppqglagqesdftapdvwypggmvsrvpypsptcvksemgpwmdsysgpygd 536

Qy    541 MRLETARDHVLPIDYFPPQKTCLICGDEASGHYGALTCGSCKVFFKRAAEGKQKYLCA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    537 mrletardhvlpidyyfppqktclicgdeasgchygaltcgsckvffkraaegkqkylca 596

Qy    601 SRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNLKLQEEGEASSTTSPTEETT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    597 srndctidkfrrkncpscrlrkcyeagmtlgarklkklgnlklqeegeassttspteett 656

Qy    661 QKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVV 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657 qkltvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvv 716

Qy    721 KWAKALPGPRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMH 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717 kwakalpgprnlhvddqmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfneyrmh 776

Qy    781 KSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIK 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    777 ksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipvdglknqkffdelrmnyik 836
```

```
Qy    841 ELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    837 eldriiackrknptscsrrfyqltklldsvqpiarelhqftfdllikshmvsvdfpemma 896

Qy    901 EIISVQVPKILSGKVKPIYFHTQ 923
          ||||||||||||||||||||||
Db    897 eiisvqvpkilsgkvkpiyfhtq 919

RESULT   2
AAY78914
ID   AAY78914 standard; protein; 919 AA.
XX
AC   AAY78914;
XX
DT   23-MAY-2000  (first entry)
XX
DE   Human androgen receptor (AR) amino acid sequence.
XX
KW   Androgen receptor; AR; androgen-independent activation; inhibitor;
KW   cancer; benign prostatic hyperplasia; hirsutism; androgenic alopecia;
KW   acne; breast cancer; Kennedy disease; prostate cancer.
XX
OS   Homo sapiens.
XX
PN   WO20000001813-A2.
XX
PD   13-JAN-2000.
XX
PF   30-JUN-1999;  99WO-CA00604.
XX
PR   30-JUN-1998;  98US-0091871.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI   Sadar MD,  Bruchovsky N,  Gout PW,  Snoek R,  Mawji NR;
XX
DR   WPI; 2000-182113/16.
XX
PT   Novel non-androgen ligand binding peptides for inhibiting
PT   androgen-independent activation of androgen receptor, used for
PT   screening compounds and for treatment of androgen-mediated diseases
PT   such as prostate cancer -
XX
PS   Disclosure; Page 7; 32pp; English.
XX
CC   This sequence represents the human androgen receptor (AR) amino acid
CC   sequence. The invention relates to a fragment of the AR corresponding to
CC   amino acids 234-391 (see AAY78913). The fragment mediates
CC   androgen-independent activation of the AR. The androgen receptor acts as
CC   a transcription factor, regulating the expression of certain
CC   androgen-responsive genes. Interaction of the AR with the protein kinase
CC   A signal transduction pathway involves interaction with the androgen
CC   independent region. The AR fragment and peptides derived from it can be
CC   used as agents for inhibiting androgen independent activation of the
CC   androgen receptor, as activation domains, and as a tool for screening
CC   for compounds which affect androgen-independent activation of the AR.
CC   The peptides, when used in combination with androgen deprivation,
CC   effectively limit androgen mediated disease progression. These diseases
CC   include cancer, benign prostatic hyperplasia, hirsutism, androgenic
CC   alopecia, acne, breast cancer, Kennedy disease, and especially prostate
CC   cancer. The peptides and nucleic acids encoding them, are especially used
CC   for the treatment of androgen-mediated diseases, especially prostate
CC   tumours in patients deprived of androgen.
XX
SQ   Sequence    919 AA;

Query Match             99.3%;    Score 4880;  DB 21;  Length 919;
Best Local Similarity   99.6%;    Pred. NO. 0;
Matches   919;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy      1 MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGPRHPEAASAAPPGASLLLLQQQ 60
```

Fri Jan  4 12:39:38 2002

```
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1  mevqlglgrvyrppsktyrgafqnlfqsvreviqnpgprhpeaasaappgaslll--   57
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    61  QQQQQQQQQQQQQQQQQQQQQQQETSPRQQQQQQGEDGSPQAHRRGPTGYLVLDEEQQPSQ  120
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    58  -qqqqqqqqqqqqqqqqqqqqqqetsprqqqqqqgedgspqahrrgptgylvldeeqqpsq  116
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  PQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSCS  180
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   117  pqsalechpergcvpepgaavaaskglpqqlpappdeddsaapstlsllgptfpglsscs  176
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181  ADLKDILSEASTMQLLQQQQQEAVSEGSSSGRAREASGAPTSSKDNYLGGTSTISDNAKE  240
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   177  adlkdilseastmqllqqqqqeavsegsssgrareasgaptsskdnylggtstisdnake  236
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241  LCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLLDDS  300
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   237  lckavsvsmglgvealehlspgeqlrgdcmyapllgvppavrptpcaplaeckgslldds  296
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301  AGKSTEDTAEYSPFKGGYTKGLEGESLGCSGSAAAGSSGTLELPSTLSLYKSGALDEAAA  360
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   297  agkstedtaeyspfkggytkglegeslgcsgsaaagssgtlelpstlslyksgaldeaaa  356
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361  YQSRDYYNFPLALAGPPPPPPPPHPHARIKLENPLDYGSAWAAAAAQCRYGDLASLHGAG  420
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   357  yqsrdyynfplalagppppppppphpharikenpldygsawaaaaaqcrygdlaslhgag  416
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421  AAGPGSGSPSAAASSSWHTLFTAEEGQLYGPCGGGGGGGGGGGGGGGGGGGGGGGGGEAGA  480
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   417  aagpgsgspsaaassswhtlftaeeeglygpcgggggggggggggggggggggggggeaga  476
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481  VAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTCVKSEMGPWMDSYSGPYGD  540
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   477  vapygytrppqglagqesdftapdvwypggmvsrvpypsptcvksemgpwmdsysgpyyd  536
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541  MRLETARDHVLPIDYYFPPQQKTCLICGDEASGCHYGALTCGSCKVFFKRAAEGKQKYLCA  600
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   537  mrletardhvlpidyyfppqktclicgdeasgchygaltcgsckvffkraaegkqkylca  596
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601  SRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNLKLQEEGEASSTTSPTEETT  660
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   597  srndctidkfrrkncpscrlrkcyeagmtlgarklkklgnlklqeegeassttspteett  656
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661  QKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVV  720
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   657  qkltvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvv  716
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721  KWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMH  780
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   717  kwakalpgfrnlhvddqmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfneyrmh  776
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781  KSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIK  840
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   777  ksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipvdglknqkffdelrmnyik  836
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841  ELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMA  900
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   837  eldriiackrknptscsrrfyqltklldsvqpiarelhqftfdlllikshmvsvdfpemma  896
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901  EIISVQVPKILSGKVKPIYFHTQ  923
Db          |||||||||||||||||||||||
Qy   897  eiisvqvpkilsgkvkpiyfhtq  919
```

RESULT     3
AAP93109
ID   AAP93109 standard; protein; 919 AA.
XX
AC   AAP93109;
XX
DT   19-MAR-1990  (first entry)
XX
DE   Human androgen receptor.

---

```
XX
KW   Human androgen receptor; ployclonal antibody; cancer.
XX
OS   Homo sapiens.
XX
PN   WO809791-A.
XX
PD   19-OCT-1989.
XX
PF   13-APR-1989;   89WO-US01548.
XX
PR   14-APR-1988;   88US-0182646.
XX
PA   (UYNC-) UNIVERSITY OF NORTH CAROLINA.
XX
PI   French FS, Wilson EM, Joseph DR, Lubahn DB;
XX
DR   WPI; 1989-324206/44.
DR   N-PSDB; AAN91772.
XX
PT   DNA encoding androgen receptor protein - useful for transforming
PT   eukaryotic hosts for protein expression and subsequent antibody prodn.
XX
PS   Disclosure; Fig. 4; 41pp; English.
XX
CC   Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC   antibodies.  These are used for the detection and quantification of AR in
CC   the presence of endogenous androgen, as androgen will not interfere with
CC   binding.  They may be used in assays to determine and quantify cellular
CC   distribution of AR in tumour tissue, and are esp. useful for evaluating
CC   prostate cancers to determine responsiveness to androgen withdrawal
CC   therapy.
XX
SQ   Sequence  919 AA;

Query Match        99.2%;   Score 4872;   DB 10;   Length 919;
Best Local Similarity 99.5%;   Pred. No. 0;
Matches 918;  Conservative  0;  Mismatches  0;  Indels  4;  Gaps  1;

Qy     1  MEVQLGLGRVYRPPSKTYRGAFQNLFQSVREVIQNPGPRHPEAASAAPPGASLLLQQQ   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| --
Db     1  mevqlglgrvyrppsktyrgafqnlfqsvreviqnpgprhpeaasaappgaslll---  67
Qy    61  QQQQQQQQQQQQQQQQQQQETSPRQQQQQQGEDGSPQAHRRGPTGYLVLDEEQQPSQ  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    58  -qqqqqqqqqqqqqqqqqqqqetsprqqqqqqgedgspqahrrgptgylvldeeqqpsq  116
Qy   121  PQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSCS  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   117  pqsalechpergcvpepgaavaaskglpqqlpappdeddsaapstlsllgptfpglsscs  176
Qy   181  ADLKDILSEASTMQLLQQQQQEAVSEGSSSGRAREASGAPTSSKDNYLGGTSTISDNAKE  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   177  adlkdilseastmqllqqqqqeavsegsssgrareasgaptsskdnylggtstisdnake  236
Qy   241  LCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLLDDS  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237  lckavsvsmglgvealehlspgeqlrgdcmyapllgvppavrptpcaplaeckgslldds  296
Qy   301  AGKSTEDTAEYSPFKGGYTKGLEGESLGCSGSAAAGSSGTLELPSTLSLYKSGALDEAAA  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   297  agkstedtaeyspfkggytkglegeslgcsgsaaagssgtlelpstlslyksgaldeaaa  356
Qy   361  YQSRDYYNFPLALAGPPPPPPPPHPHARIKLENPLDYGSAWAAAAAQCRYGDLASLHGAG  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   357  yqsrdyynfplalagppppppppphpharikenpldygsawaaaaaqcrygdlaslhgag  416
Qy   421  AAGPGSGSPSAAASSSWHTLFTAEEGQLYGPCGGGGGGGGGGGGGGGGGGGGGGGGEAGA  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   417  aagpgsgspsaaassswhtlftaeeeglygpcggggggggggggggggggggggggeaga  476
Qy   481  VAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTCVKSEMGPWMDSYSGPYGD  540
```

CC     dependence polypeptides consisting of p75NTR, androgen receptor, DCC,
CC     huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2,
CC     SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable
CC     of inducing cell death and can be used to develop products to mediate or
CC     inhibit apoptosis. The methods can be used for reducing the severity of
CC     a proapoptotic dependence domain mediated pathological conditions e.g.
CC     Huntington's disease, Alzheimer's disease, Kennedy's disease,
CC     Spinocerebellar ataxia, dentatorubropallidoluysian atrophy,
CC     Machado-Joseph disease, stroke or head trauma. They can also be used for
CC     reducing the severity of a pathological condition mediated by upregulated
CC     cell proliferation or cell survival e.g. neoplastic, malignant,
CC     autoimmune or fibrotic conditions. This sequence represents a human
XX     androgen receptor described in the method of the invention.
SQ     Sequence   918 AA;

Query Match             98.5%;     Score 4838.5;   DB 20;   Length 918;
Best Local Similarity   98.7%;     Pred. No. 6.4e-318;
Matches  914;   Conservative    0;   Mismatches    0;   Indels    11;   Gaps    2;

Qy      1  MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGPRHPEAASAAPPGASLLLLQQQ  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  mevqlglgrvyprppsktyrgafqnlfqsvreviqnpgprhpeaasaappgaslll---  57

Qy     61  QQQQQQQQQQQQQQQQQQETSPRQQQQQQGEDGSPQAHRRGPTGYLVLDEEQQPSQ  120
           |||||||||||||||||||||||||||||||||||||
Db     58  ----qqggqgqgqgqgqgqgetsprqgqqqgqgedgspqahrrgptgylvldeeqqpsq  112

Qy    121  PQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSCS  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    113  pqsalechpergcvpepgaavaaskglpqqlpappdeddsaapstlsllgptfpglsscs  172

Qy    181  ADLKDILSEASTMQLLQQQQEAVSEGSSSGRAREASCAPTSSKDNYLGGTSTISDNAKE  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173  adlkdilseastmqllqqqqeavsegsssgrareasgaptsskdnylggtstisdnake  232

Qy    241  LCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLLDDS  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    233  lckavsvsmglgvealehlspgeqlrgdcmyapllgvppavrptpcaplaeckgslldds  292

Qy    301  AGKSTEDTAEYSPFKGGYTKGLEGESLGCGSAAAGSSGTLELPSTLSLYKSGALDEAAA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    293  agkstedtaeyspfkggytkglegeslgcsgsaaagssgtlelpstlslyksgaldeaaa  352

Qy    361  YQSRDYNFPLALAGPPPPPPPPPHARIKLENPLDYGSAWAAAAAQCRYGDLASLHGAG  420
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    353  yqsrdynfplalagpppppppphariklenpldyggsawaaaaqcrygdlaslhgag  412

Qy    421  AAGPGSGSPSAAASSSWHTLFTAEEGQLYGPC--GGGGGGGGGGGGGGGGGGGGGGGGE  477
           ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db    413  aagpgsgspsaaassswhtlftaeeqglyapcgggggggggggggggggggggggggge  472

Qy    478  AGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTCVKSEMGPWMDSYSGP  537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473  aeavapygytrppqglagqesdftapdvwypggmvsrvpypsptcvksemgpwmdsysgp  532

Qy    538  YGDMRLETARDHVLPIDYYFPPQKTCLICGDEASGCHYGALTCGSCKVFFKRRAAEGKQKY  597
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    533  ygdmrletardhvlpidyyfppqktclicgdeasgchygaltcgsckvffkrraaegkqky  592

Qy    598  LCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNLKLQEEGEASSTTSPTE  657
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    593  lcasrndctidkfrrkncpscrlrkcyeagmtlgarklkklgnlklqeegeassttspte  652

Qy    658  ETTQKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLV  717
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    653  ettqkltvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlv  712

Qy    718  HVVKWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEY  777
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    713  hvvkwakalpgfrnlhvddqmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfney  772

---

Db    477  vapygytrppqglagqesdftapdvwypggmvsrvpypsptcvksemgpwmdsysqpygd  536
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541  MRLETARDHVLPIDYYFPPQKTCLICGDEASGCHYGALTCGSCKVFFKRRAAEGKQKYLCA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    537  mrletardhvlpidyyfppqktclicgdeasgchygaltcgsckvffkrraaegkqkylca  596

Qy    601  SRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNLKLQEEGEASSTTSPTEETT  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    597  srndctidkfrrkncpscrlrkcyeagmtlgarklkklgnlklqeegeasotttpteett  656

Qy    661  QKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVV  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657  qkltvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvv  716

Qy    721  KWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMH  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717  kwakalpgfrnlhvddqmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfneyrmh  776

Qy    781  KSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIK  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    777  ksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipvdglknqkffdelrmnyik  836

Qy    841  ELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFFPEMMA  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    837  eldriiackrknptscsrrfyqltklldsvqpiarelhqftfdlllikshmvsvdffpemma  896

Qy    901  EIISVQVPKILSGKVKPIYFHTQ  923
           |||||||||||||||||||||||
Db    897  eiisvqvpkilsgkvkpiyfhtq  919

RESULT    4
AAY33491
ID     AAY33491  standard;  Protein;  918 AA.
XX
AC     AAY33491;
XX
DT     19-JAN-2000  (first entry)
XX
DE     Human androgen receptor protein.
XX
KW     Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;
KW     huntingtin polypeptide; Machado-Joseph disease; SCA1; SCA2; SCA6;
KW     atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
KW     Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
KW     dentatorubropallidoluysian atrophy; cell proliferation; cell survival;
KW     neoplastic; malignant; autoimmune; fibrotic.
XX
OS     Homo sapiens.
XX
PN     WO9945944-A1.
XX
PD     16-SEP-1999.
XX
PF     11-MAR-1999;  99WO-US05250.
PR     12-MAR-1998;  98US-0041886.
XX
PA     (BURN-)  BURNHAM INST.
XX
PI     Bredesen DE,  Rabizadeh S;
XX
DR     WPI; 1999-561617/47.
DR     N-PSDB;  AAZ23424.
XX
PT     New proapoptotic dependence peptides, used to develop products for
PT     treating, e.g. Alzheimer's disease -
XX
PS     Disclosure; Page 90-93; 199pp; English.
XX
CC     This invention describes novel pure proapoptotic dependence peptides
CC.    which comprise a sequence of an active dependence domain selected from

```
Qy   778 RMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMN 837
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   773 rmhksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipvdglknqkffdelrmn 832

Qy   838 YIKELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPE 897
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   833 yikeldriiackrknptscsrrfyqltklldsvqpiarelhqftfdlliikshmvsvdfpe 892

Qy   898 MMAEIISVQVPKILSGKVKPIYFHTQ 923
         ||||||||||||||||||||||||||
Db   893 mmaeiisvqvpkilsgkvkpiyfhtq 918
```

RESULT 5
AAR12223
ID   AAR12223 standard; Protein; 918 AA.
XX
AC   AAR12223;
XX
DT   20-AUG-1991 (first entry)
XX
DE   Human androgen receptor.
XX
KW   hAR; DNA-binding protein; steroid hormone.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   Domain         556..626
FT                  /label= DNA-binding domain
FT                  /note= "cysteine-rich"
XX
PN   WO9107423-A.
XX
PD   30-MAY-1991.
XX
PF   19-OCT-1990; 90WO-US06015.
XX
PR   17-NOV-1989; 89US-0438775.
XX
PA   (ARCH-) ARCH DEV CORP.
XX
PI   Liao S, Chang C;
XX
DR   WPI; 1991-178048/24.
DR   N-PSDB; AAQ12001.
XX
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
XX
PS   Claim 25; Fig 3; 79pp; English.
XX
CC   This sequence was deduced from a cDNA clone isolated by screening
CC   commercially available human testis and prostate lambda gt11 cDNA
CC   libraries. The sequence is very similar to that of rat cDNA and in
CC   the DNA-binding domain it is identical to that of rAR DNA-binding
CC   domain. Homology comparisons with other known steroid receptors
CC   indicate that hAR is more closely related to glucocorticoid,
CC   mineralo-corticoid and progesterone receptors than to v-erb-A or to
CC   receptors for oestrogen, vitamin D and thyroid hormones.
XX
SQ   Sequence  918 AA;

Query Match      98.3%;  Score 4827.5; DB 12; Length 918;
Best Local Similarity 98.5%;  Pred. No. 3.4e-317;
Matches 912;  Conservative 1;  Mismatches 2;  Indels 11;  Gaps 2;

Qy   1 MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGPRHPEAASAAPPGASLLLQQQ 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 mevqlglgrvyprppsktyrgafqnlfqsvreviqnpgpprhpeaasaappgasll-- 57
Qy  61 QQQQQQQQQQQQQQQQQQQQQQQQETSPRQQQQQQGEDGSPQAHRRGPTGYLVLDEEQQPSQ 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  58 ----qqgqqqqqqqqqgqqgetsprqqqqqqgedgspqahrrgptgylvldeeqqpsq 112

Qy 121 PQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSCS 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 113 pqsalechpergcvpepgaavaaskglpqqlpappdeddsaapstlsllgptfpglsscs 172

Qy 181 ADLKDILSEASTMQLLQQQQQEAVSEGSSSGRAREASGAPTSSKDNYLGGTSTISDNAKE 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 173 adlkdilseastmqllqqqqqeavsegsssgrareasgaptsskdnylggtstisdnake 232

Qy 241 LCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLLDDS 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 233 lckavsvsmglgvealehlspgeqlrgdcmyapllgvppavrptpcaplaeckgslldds 292

Qy 301 AGKSTEDTAEYSPFKGGYTKGLEGESLGCSGSAAAGSSGTLELPSTLSLYKSGALDEAAA 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 293 agkstedtaeyspfkggytkglegeslgcsgsaaagssgtlelpstlslyksgaldeaaa 352

Qy 361 YQSRDYNFPLALAGPPPPPPPPHPHARIKLENPLDYGSAWAAAAQCRYGDLASLHGAG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 353 yqsrdyynfplalagppppppppphpariklenpldygsawaaaaaqcryydiaslhgag 412

Qy 421 AAGPGSGSPSAAASSWHTLFTAEEQLYGPC-1-GGGGGGGGGGGGGGGGGGGE 477
        ||||||||||||||||||||||||||||||      ||||||||||||||||||
Db 413 aarpgsgspsaaassswhtlftaeeeglygpcgcgggggggggggggggggggge 472

Qy 478 AGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTCVKSEMGPWMDSYSGP 537
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 473 aeavapygytrppqglagqesdftapdvwypggmvsrvpypsptcvksemgpwmdsysgp 532

Qy 538 YGDMRLETARDHVLPIDYFPPQKTCLICGDEASGCHYGALTCGSCKVFFKRAAEGKQKY 597
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 533 ygdmrletardhvlpidyyfppqktclicgdeasgchygaltcgsckvffkraaegkqky 592

Qy 598 LCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKLGNLKLQEEGEASSTTSPTE 657
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 593 lcasrndctidkfrrkncpscrlrkcyeagmtlgarklklgnlklqeegeassttspte 652

Qy 658 ETTQKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLV 717
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 653 ettqkltvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlv 712

Qy 718 HVVKWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEY 777
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 713 hvvkwakalpgfrnlhvddqmavieyswmglmvfamgwrsftnvnsrmlyfapdlvfney 772

Qy 778 RMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMN 837
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 773 rmhksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipvdglknqkffdelrmn 832

Qy 838 YIKELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPE 897
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 833 yikeldriiackrknptscsrrfyqltklldsvqpiarelhqftfdlliikshmvsvdfpe 892

Qy 898 MMAEIISVQVPKILSGKVKPIYFHTQ 923
        ||||||||||||||||||||||||||
Db 893 mmaeiisvqvpkilsgkvkpiyfhtq 918
```

RESULT 6
AAP90996
ID   AAP90996 standard; protein; 919 AA.
XX
AC   AAP90996;
XX
DT   28-FEB-1990 (first entry)
XX
DE   Human androgen receptor DNA clone.
XX
KW   Androgen receptor; TR2 polypeptide;
XX
OS   Homo sapiens.

```
XX
FH   Key             Location/Qualifiers
FT   Region          1..919
FT                   /tag= a
FT                   /product=98 kD polypeptide
FT   Region          185..919
FT                   /tag= b
FT                   /product=79 kD polypeptide
XX
PN   WO8909223-A.
XX
RD   05-OCT-1989.
XX
PF   24-MAR-1989;    89WO-US01238.
XX
PR   30-MAR-1988;    88US-0176107.
XX
PA   (ARCH.) ARCH DEVELOPMENT CORP.
XX
PI   Liao S, Chang C;
XX
DR   WPI; 1989-309501/42.
DR   N-PSDB; AAN91577.
XX
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assays and for
PT   purification.
XX
PS   Claim 8; Fig 3; 60pp; English.
XX
CC   The protein is used to raise antibodies for receptor assays and
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon; the 79 kD product
CC   starts from the second.
XX
SQ   Sequence 919 AA;

Query Match             98.0%;
Best Local Similarity   98.4%;
Matches   912;   Conservative   0;   Mismatches   0;   Indels   3;   Gaps   3;

Qy    1 MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGPRHPEAASAAPPGASLLLLQQQQ 60
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    1 mevqlglgrvyprpppsktyrgafqnlfqsvreviqnpgprhpeaasaappgaslll---- 57

Qy   61 QQQQQQQQQQQQQETSPRQQQQQ-GEDGSPQAHRRGPTGYLVLDEEQQPS 119
          ||||||||||||| ||||||| |||||||||||||||||||||||||||
Db   58 ----qqqqqqqqqqqetpqqqqqhgedgspqahrrgptgylvldeeqqps 112

Qy  120 QPQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSC 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  113 qpqsalechpergcvpepgaavaaskglpqqlpappdeddsaapstlsllgptfpglssc 172

Qy  180 SADLKDILSEASTMQLLQQQQQQEAVSEGSSSGRAREASGAPTSSKDNYLGGTSTISDNAK 239
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  173 sadlkdilseastmqllqqqqqqeavsegsssgrareasgaptsskdnylggtstisdnak 232

Qy  240 ELCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLLDD 299
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  233 elckavsvsmglgvealehlspgeqlrgdcmyapllgvppavrptpcaplaeckgslidd 292

Qy  300 SAGKSTEDTAEYSPFKGGYTKGLEGESLGCSSAAAGSSGTLELPSTLSLYKSGALDEAA 359
          |||||||||||||||||||||||||| ||| | ||||||||||||||||||||||||
Db  293 sagkstedtaeyspfkggytkgleresigcsgsaaagssgtlelpstlslyksgaldeaa 352

Qy  360 AYQSRDYYNFPLALAGPPPPPPPHPHARIKLENPLDYGSAWAAAAQCRYGDLASLHGA 419
          |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db  353 ayqsrdyynfplalagppppppphpharikienpldygsawaaaaaqcrygdlaslhga 412

Qy  420 GAAGPGSGSPSAAAASSWHTLFTAEEGQLYGPC---GGGGGGGGGGGGGGGGGGGGGGG 476
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  413 gaagpgsgspsaaaasswhtlftaeegqlygpcggggggggggggggggggggggggggg 472
```

```
Qy  477 EAGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTCVKSEMGPWMDSYSG 536
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  473 eaeavapygytrppqglaggqesdftapdvwypggmvsrvpypsptcvksemgpwmdsysg 532

Qy  537 PYGDMRLETARDHVLPIDYYFPPQKTCLICGDEASGCHYGALTCGSCKVFFKRAAEGKQK 596
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  533 pygdmrletardhvlpidyyfppqktclicgdeasgchygaltcgsckvffkraaaegkqk 592

Qy  597 YLCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNLKLQEEGEASSTTSPT 656
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  593 ylcasrndctidkfrrkncpscrlrkcyeagmtlgarklklgnlklqeegeasssttspt 652

Qy  657 EETTQKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQL 716
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  653 eettqkltvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfaallssinelgerql 712

Qy  717 VHVVKAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNE 776
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  713 vhvvkwakalpgfrnlhvddqmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfne 772

Qy  777 YRMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRM 836
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  773 yrmhksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipvdglknqkffdelrm 832

Qy  837 NYIKELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELIQFTFDLLIKSHMVSVDFP 896
          ||||||||||||||||||||||||||||||||||||||||| |||||| |||||||| |
Db  833 nyikeldriiackrknptscsrrfyqltklldsvqpiarelhqftfdlliikshmvsvdfp 892

Qy  897 EMMAEIISVQVPKILSGKVKPIYFHTQ 923
          |||||||||||||||||||||||||||
Db  893 emmaeiisvqvpkilsgkvkpiyfhtq 919
```

```
RESULT 7
AAP91006
ID   AAP91006 standard; protein; 902 AA.
XX
AC   AAP91006;
XX
DT   28-FEB-1990 (first entry)
XX
DE   Rat androgen receptor DNA clone.
XX
KW   Androgen receptor; TR2 polypeptide;
XX
OS   Rat.
XX
FH   Key             Location/Qualifiers
FT   Region          1..902
FT                   /tag= a
FT                   /product=98 kD polypeptide
FT   Region          170..902
FT                   /tag= b
FT                   /product=79 kD polypeptide
XX
PN   WO8909223-A.
XX
RD   05-OCT-1989.
XX
PF   24-MAR-1989;    89WO-US01238.
XX
PR   30-MAR-1988;    88US-0176107.
XX
PA   (ARCH.) ARCH DEVELOPMENT CORP.
XX
PI   Liao S, Chang C;
XX
DR   WPI; 1989-309501/42.
DR   N-PSDB; AAN91578.
XX
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
```

```
PT   purification.
XX
PS   Claim 8; Fig 3; 60pp; English.
XX
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon; the 79 kD product
CC   starts from the second.
XX
SQ   Sequence   902 AA;

Query Match
Best Local Similarity  85.5%;            Score 4201.5;  DB 10;  Length 902;
                       84.6%;  Pred. No. 5.6e-275;
Matches 796; Conservative 36; Mismatches 52; Indels 57; Gaps 4;

Qy    1  MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGRHPEAASAAPPGASLLLQQQ    60
         |||||||||||||||||||||||||||||||:||||||||| ||||||| ||||
Db    1  mevhlglgrvyprppsktyrgafqnlfqsvreaiqnpgprhpeaasiappgacl------  54

Qy   61  QQQQQQQQQQQQQQQQQQQQQQQETSPRQQQQQQ-GEDGSPQAHRRGPTGYLVLDEEQQPS 119
                                   :|:||||||| |||||||| ::|||:|||||||||
Db   55  --------------qqrgetsprrrrqqhpedgspqahirgttgylaleeeqqps        96

Qy  120  QPQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSC 179
         ||:||| ||| |||||||||||  ||:||||  ||||::|||||||||||||:||||||||
Db   97  qqgsaseqhpesgclpepgaatapgkglpqgppapdqddsaapstlsligptfpglssc   156

Qy  180  SADLKDILSEASTMQLL-----------...QQQQEAVSEGSSSGRAREASGAPTS     222
         ||:|||||||:||||||                |||:|||:|||:||||| ||||:
Db  157  sadikdilseagtmqllqqqqqqqqqqqqqqqqqqqqqevisegsssvrareatgapss  216

Qy  223  SKDNYLGGTSTISDNAKELCKAVSVSMGLGVEALEHLSPEQLRGDCMYAPLLGVPPAVR  282
         |||:||| :|||||:|||||||| ||||||||||||||||:||||||||| ||||||||
Db  217  skdsylggnstisdsakelckavssvsmglgvealehlspgeqlrgdcmyasllgvppavr 276

Qy  283  PTPCAPLAECKGSLLLDSAGKSTEDTAEYSPFKGGYTKGLEGESLGCSGSAAAGSSGTLE  342
         |||||||||||||| ||:  |||| ||||| ||||| |||||||||||||   ||||||
Db  277  ptpcaplaeckgslsldegpgkgteetaeyssfkggyakglegeslgcsgsseagssgtle 336

Qy  343  LPSTLSLYKSGALDEAAAYQSRDYYNFPLALAGPPPPPPPPHPHARIKLENPLDYGSAWA  402
         :||||||||||:|||||| | ||||||||||:||| ||||:|||||||||||||||:||
Db  337  ipsslslyksgavdeaaaaynrdyynfplalsgpphpppppthphariklenpldygsawa 396

Qy  403  AAAAQCRYGDLASLHGAGAGPGSGSPSAAASSSWHTLFTAEEGQLYGPCGGGGGGGGGG  462
         |||||||||||||||| : |||| |||| |||||||||||||:||||
Db  397  aaaaqcrygdlaslhggsvagpstgsppatassswhtlftaeeeglygp-----------  445

Qy  463  GGGGGGGGGGGEAGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTC    522
              ||||| |||||||||||||| |:|||| |:::|||::||||::||||||||:
Db  446  ----gggggsspsdagpvapygytrppqglasqegdfsasevwypggvvnrvpypspsc   501

Qy  523  VKSEMGPWMDSYSGPYGDMRLETARDHVLPIDYFPPQKTICLICGDEASGCHYGALTCGS 582
         |||||||||: ||||||||||: :|||||||||:|||||:|||||||||||||||||||||
Db  502  vksemgpwmenysgpygdmrldstrdhvlpidyyfppqktclicgdeasgchygaltcgs  561

Qy  583  CKVFFKRAAEGKQKYLCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNLK 642
         |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  562  ckvffkraaegkqkylcasrndctidkfrrknpscrlrkcyeagmtlgarklkklgnlk   621

Qy  643  LQEEGEASSTTSPTEETTQKLTVSHIEGYECOPIFLNVLEAIEPGVVCAGHDNNQPDSFA 702
         |||||: ||:||| :  |||||||||||||||||||||||||||||||||||||||||||
Db  622  lqeegenssagsptedpsqkmtvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfa  681

Qy  703  ALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVN 762
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  682  allsslnelgerqlvhvvkwakalpgfrnlhvddqmaviqyswmglmvfamgwrsftnvn  741

Qy  763  SRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPV 822
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  742  srmlyfapdlvfneyrmhksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipv  801

Qy  823  DGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSSRRFYQLTKLLDSVQPIARELHQFTF 882
         |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  802  dglknqkffdelrmnyikeldriiackrknptscssrrfyqltklldsvqplarelhqftf 861

Qy  883  DLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIYFHTQ  923
         ||||||| |||||||||||||||||||||||||||||||||
Db  862  dllikshnvsvdfpemmaeiisvqvpkilsgkvkpiyfhtq  902
```

```
RESULT 8
AAP93110
ID   AAP93110 standard; protein; 902 AA.
AC   AAP93110;
XX
DT   19-MAR-1990  (first entry)
XX
DE   Rat androgen receptor.
XX
KW   Rat androgen receptor; monoclonal antibody; ployclonal antibody; cancer.
XX
OS   Rattus rattus.
XX
PN   WO8909791-A.
XX
PD   19-OCT-1989.
XX
PF   13-APR-1989;  89WO-US01548.
XX
PR   14-APR-1988;  88US-0182646.
XX
PA   (UYNC) UNIVERSITY OF NORTH CAROLINA.
XX
PI   French FS, Wilson EM, Joseph DR, Lubahn DB;
XX
DR   WPI: 1989-324206/44.
DR   N-PSDB; AAN91773.
XX
PT   DNA encoding androgen receptor protein - useful for transforming
PT   eukaryotic hosts for protein expression and subsequent antibody prodn.
XX
PS   Disclosure; Fig. 5; 41pp; English.
XX
CC   Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC   antibodies. These are used for the detection and quantification of AR in
CC   the presence of endogenous androgen, as androgen will not interfere with
CC   binding. They may be used in assays to determine and quantify cellular
CC   distribution of AR in tumour tissue, and are esp. useful for evaluating
CC   prostate cancers to determine responsiveness to androgen withdrawal
CC   therapy.
XX
SQ   Sequence   902 AA;

Query Match
Best Local Similarity  85.5%;            Score 4200.5;  DB 10;  Length 902;
                       84.6%;  Pred. No. 6.5e-275;
Matches 796; Conservative 36; Mismatches 52; Indels 57; Gaps 4;

Qy    1  MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGRHPEAASAAPPGASLLLQQQ    60
         |||||||||||||||||||||||||||||||:||||||||| ||||||| ||||
Db    1  mevhlglgrvyprppsktyrgafqnlfqsvreaiqnpgprhpeaasiappgacl------  54

Qy   61  QQQQQQQQQQQQQQQQQQQQQQQETSPRQQQQQQ-GEDGSPQAHRRGPTGYLVLDEEQQPS 119
                                   :|:||||||| |||||||| ::|||:|||||||||
Db   55  --------------qqrgetsprrrrqqhpedgspqahirgttgylaleeeqqps        96

Qy  120  QPQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSC 179
         ||:||| ||| |||||||||||  ||:||||  ||||::|||||||||||||:||||||||
Db   97  qqgsaseqhpesgclpepgaatapgkglpqgppapdqddsaapstlsligptfpglssc   156

Qy  180  SADLKDILSEASTMQLL  222
         ||:|||||||:||||||
Db  157  sadikdilseagtmqll  216
```

PT  Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT  and antibodies for detection and quantification methods
XX
PS  Claim 25; Fig 3; 79pp; English.
XX
CC  This sequence was deduced from a cDNA clone isolated by screening
CC  a rat ventral prostate lambda gt11 library in E.coli Y1090. The
CC  sequence is very similar to that of human AR and in
CC  the DNA-binding domain it is identical to that of hAR DNA-binding
CC  domain. Homology comparisons with other known steroid receptors
CC  indicate that rAR is more closely related to glucocorticoid,
CC  mineralo-corticoid and progesterone receptors than to v-erb-A or to
CC  receptors for oestrogen, vitamin D and thyroid hormones.
XX
SQ  Sequence  902 AA;

```
Query Match       85.3%;     Score 4187.5;   DB 12;   Length 902;
Best Local Similarity   84.4%;   Pred. No. 4.9e-274;
Matches  794;  Conservative  37;  Mismatches  53;  Indels  57;  Gaps   4;

Qy    1 MEVOLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGPRHPEAASAAPPGASLLLLQQQ 60
          ||||||||||||||||||||||||||||||||||||:||||||||||:|
Db    1 mevglglgrvyprppsktyrgafqnlfqsvreaiqnpgprhpeaasiappgacl------ 54

Qy   61 QQQQQQQQQQQQQQQQQQQQQQQETSPROQQQQQ-GEDGSPQAHRRGPTGYLVLDEEQQPS 119
          ||::::||:||| :||||
Db   55 --------------qrgetsprrrrqqhpedgspqahirgtgylaleeeqqps 96

Qy  120 QPQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFPGLSSC 179
          ||| |||||||:|| ||||||  ::| ||||||| |:||||||||||||:||||||||
Db   97 qqgsaseghpesgclpepgaatapgkglpqqppapdqddsaapstlsligptfpglssc 156

Qy  180 SADLKDILSEASTMQLL------------------------------QQQQEAVSEGSSSGCRAREASGAPTS 222
          ||| ||||||:|||||                              ::|||
Db  157 sadikdilseagtmqllqqqqqqqqqqqqqqqqqqqqqeviegssvsvraraetgapss 216

Qy  223 SKDNYLGGTSTISDNAKELCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVR 282
          |||:||| ||||| ||||||||||||||||||||||||||||||||||||:||| ||||
Db  217 skdsylggnstisdsakelckavsvsmglgvealehlspgeqlrgdcmyasllggppavr 276

Qy  283 PTPCAPLAECKGSLLDDSAGKSTEDTAEYSPFKGGYTKGLEGESLGCSGSAAAGSSGTLE 342
          |||||||||||:: ||| |||||||||:||||||||||||||||||:|||||||
Db  277 ptpcaplaeckgisldegpgkgteetaeyssfkggyakglegeslgcsgsseagssgtle 336

Qy  343 LPSTLSLYKSGALDEAAAYQSRDYYNFPLALAGPPPPPPHPHARIKLENPLDYGSAWA 402
          :||||||||||:|||||||::|||||||||||:||| |||:||||||||||:||||||
Db  337 ipsslslyksgavdeaaayqnrdyynfplalsgpphpppthphariklenpdygsawa 396

Qy  403 AAAAQCRYGDLASLHGAGAAGPGSGSPSAAAASSWHTLFTAEEGQLYGPCGGGGGGGGGG 462
          ||||||||||||||||  || |||||||||||||||||||||||||||
Db  397 aaaaqcrygdlaslhggsvarpstgsppatassswhtlftaeegglygp----------- 445

Qy  463 GGGGGGGGGGGGGGEAGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTC 522
          ||||   |||||||||||||||||||||||:|||||:|||||:||||:|||||||||
Db  446 ---ggggggsspsdagpvapygytrppqglasqegdfsasevvypggvvnrvpypspsc 501

Qy  523 VKSEMGPWMDSYSGPYGDMRLETARDHVLPIDYYFPPQKTCLICGDEASGCHYGALTCGS 582
          |||||||| |||||||||:||::||:||||:|||||||||||||||||:||||||||||||
Db  502 vksemgpwmenysgpygdmrldstrdhvlpidyyfppqktclicgdeasgchygaltcgs 561

Qy  583 CKVFFKRAAEGKQKYLCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNLK 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Qy  223 SKDNYLGGTSTISDNAKELCKAVSVSMGLGVEALEHLSPGEQLRGDCMYAPLLGVPPAVR 282
          |||:||| ||||| |||||||||||||||||||||||||||||||||||||:||| ||||
Db  217 skdsylggnstisdsakelckavsvsmglgvealehlspgeqlrgdcmyasllggppavr 276

Qy  283 PTPCAPLAECKGSLLDDSAGKSTEDTAEYSPFKGGYTKGLEGESLGCSGSAAAGSSGTLE 342
          |||||||||||:: |||| || |||||||:||||||||||||||||||:|||||||
Db  277 ptpcaplaeckgisldegpgkgteetaeyssfkggyakglegeslgcsgsseagssgtle 336

Qy  343 LPSTLSLYKSGALDEAAAYQSRDYYNFPLALAGPPPPPPHPHARIKLENPLDYGSAWA 402
          :|||||||||||:|||||||::||||||||||:||| |||:||||||||||:||||||
Db  337 ipsslslyksgavdeaaayqnrdyynfplalsgpphpppthphariklenpdygsawa 396

Qy  403 AAAAQCRYGDLASLHGAGAAGPGSGSPSAAAASSWHTLFTAEEGQLYGPCGGGGGGGGGG 462
          ||||||||||||||||  || |||||||||||||||||||||||||||
Db  397 aaaaqcrygdlaslhggsvarpstgsppatassswhtlftaeegglygp----------- 445

Qy  463 GGGGGGGGGGGGGGEAGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTC 522
          ||||   |||||||||||||||||||||||:|||||:|||||:||||:|||||||||
Db  446 ---ggggggsspsdagpvapygytrppqglasqegdfsasevvypggvvnrvpypspsc 501

Qy  523 VKSEMGPWMDSYSGPYGDMRLETARDHVLPIDYYFPPQKTCLICGDEASGCHYGALTCGS 582
          |||||||| |||||||||:||::||:||||:|||||||||||||||||:||||||||||||
Db  502 vksemgpwmenysgpygdmrldstrdhvlpidyyfppqktclicgdeasgchygaltcgs 561

Qy  583 CKVFFKRAAEGKQKYLCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLGNLK 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  562 ckvffkraaegkqkylcasrndctidkfrrkncpscrlrkcyeagmtlgarklkklgnlk 621

Qy  643 LQEEGEASSTTSPTEETTQKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFA 702
          ||||||:|||  ||||  ||||:||||||||||||||||||||||||||||||||||||
Db  622 lqeegenssagsptedpsqkmtvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfa 681

Qy  703 ALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVN 762
          ||||:|||| |||||||||||||||||||||||||||||||||||:||||||:|||||||
Db  682 allssinelgeqlvhvvkwakalpgfrnlhvddqmaviqyswmglmvfamgwrsftnvn 741

Qy  763 SRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPV 822
          |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db  742 srmlyfapdlvfneyrmhksrmysqvrmrhlsqefgwlqitpqeflcmkalllfsiipv 801

Qy  823 DGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRREYQLTKLLDSVQPIARELHQFTF 882
          ||||||||||||||||||||||||||||||||||| ||:|||||||||||||||||||||
Db  802 dglknqkffdelrmnyikeldriiackrknptscssrrfyqltklldsvqpiarelhqftf 861

Qy  883 DLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIYFHTQ 923
          |||||||||||||||| |||||||||||||||||||||||
Db  862 dllikshmvsvdfpenmaeiisvqvpkilsgkvkpiyfhtq 902

RESULT  9

```
Db   562  ckvffkraaegkgkylcasrndctidkfrrkncpscrlrkcyeagmtlgarklklgnlk  621

Qy   643  LQEEGEASSTTSPTEETTQKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFA  702
          |||||| ||  |||| :||::||:||||||||||||||||||||||||||||||||||||
Db   622  lqeegenssagsptedpsqkmtvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfa  681

Qy   703  ALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVN  762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682  allsslnelgerqlvhvvkwakalpgfrnlhvddqmaviqyswmglmvfamgwrsftnvn  741

Qy   763  SRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPV  822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742  srmlyfapdlvfneyrmhksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipv  801

Qy   823  DGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTF  882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   802  dglknqkffdelrmnyikeldriiackrknptscsrrfyqltklldsvqpiarelqqftf  861

Qy   883  DLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIYFHTQ  923
          |||||||||||||||||||||||||||||||||||||||||
Db   862  dllikshmvsvdfpemmaeiisvqvpkilsgkvkpiyfhtq  902
```

```
RESULT 10
AAY21627
ID   AAY21627 standard; protein; 452 AA.
XX
AC   AAY21627;
XX
DT   11-AUG-1999  (first entry)
XX
DE   Ligand binding domain of nuclear receptor hAR.
XX
KW   Thyroid hormone receptor; aromatic compound; ligand binding domain;
KW   alpha-glycerophosphate dehydrogenase; cardiac; obesity; triglyceride;
KW   plasma cholesterol; anti-hypertriglyceridaemic; atherosclerosis; GPDH;
KW   thyroid hormone replacement therapy; nuclear receptor.
XX
OS   Homo sapiens.
XX
PN   WO992696-A2.
XX
PD   03-JUN-1999.
XX
PF   25-NOV-1998;  98WO-US25296.
XX
PR   26-NOV-1997;  97US-0980115.
XX
PA   (REGC-) UNIV CALIFORNIA.
XX
PI   Apriletti JW, Baxter JD, Fletterick RJ, Kushner PJ;
PI   Scanlan TS, Shiau AK, Wagner RL, West BL;
XX
DR   WPI; 1999-357810/30.
XX
PT   Modulating activity of a thyroid hormone receptor
XX
PS   Disclosure; Fig 3H-R; 447pp; English.
XX
CC   The invention relates to a method for modulating activity of a thyroid
CC   hormone receptor that comprises administration of an aromatic compound
CC   which fits spatially and preferentially into a thyroid hormone ligand
CC   binding domain. The aromatic compound (of a specified formula) can be
CC   used to increase alpha-glycerophosphate dehydrogenase (GPDH) levels, at
CC   levels which do not significantly modify cardiac GPDH levels and are
CC   indicated in the treatment of obesity. The compound also lower total
CC   plasma cholesterol and triglyceride levels and can be used as anti-
CC   hypertriglyceridaemic agents. The compound may also be used for treating
CC   atherosclerosis and may be indicated in thyroid hormone replacement
CC   therapy in patients with compromised cardiac function. Sequences
CC   AAY21621- 636 amino acid sequences of ligand binding domains of several
CC   members of the nuclear receptor superfamily.
XX
```

```
SQ   Sequence  452 AA;

Query Match       49.5%;   Score 2429;  DB 20;  Length 452;
Best Local Similarity  99.8%;   Pred. No. 7.9e-156;
Matches  451;  Conservative  1;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   472  GGGGGEAGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPSPTCVKSEMGPWM  531
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| :||
Db     1  gggggeagavapygytrppqglagqesdftapdvwypggmvsrvpypsptcvksemgpwm  60

Qy   532  DSYSGPYGDMRLETARDHVLPIDYYFPPQKTCLIGDEASQCHYGALTCGSCKVFFKRAA  591
          ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db    61  dsysgpygdmrletardhvlpidyyfppqktclicgdkasqchygaltcgsckvffkraa  120

Qy   592  EGKQKYLCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLKKLIGNLKLQEBGEASS  651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   121  egkqkylcasrndctidkfrrkncpscrlrkcyeagmtlgarklkklgnlklqeegeass  180

Qy   652  TTSPTEETTQKLTVSHIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNEL  711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ttspteettqkltvshiegyecqpiflnvleaiepgvvcaghdnnqpdsfaallsslnel  240

Qy   712  GERQLVHVVKWAKALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPD  771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  gerqlvhvvkwakalpgfrnlhvddqmaviqyswmglmvfamgwrsftnvnsrmlyfapd  300

Qy   772  LVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFF  831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  lvfneyrmhksrmysqcvrmrhlsqefgwlqitpqeflcmkalllfsiipvdglknqkff  360

Qy   832  DELRMNYIKELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMV  891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  delrmnyikeldriiackrknptscsrrfyqltklldsvqpiarelhqftfdllikshmv  420

Qy   892  SVDFPEMMAEIISVQVPKILSGKVKPIYFHTQ  923
          ||||||||||||||||||||||||||||||||
Db   421  svdfpemmaeiisvqvpkilsgkvkpiyfhtq  452
```

```
RESULT 11
AAR12230
ID   AAR12230 standard; protein; 630 AA.
XX
AC   AAR12230;
XX
DT   20-AUG-1991  (first entry)
XX
DE   TrpE/androgen receptor DNA-binding domain fusion protein.
XX
KW   androgen receptor; AR; DNA-binding protein; steroid hormone.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   Protein        1..323
FT                  /label= 33kD trpE protein
FT   Region         324..334
FT                  /label= 11 amino acid linker
FT   Protein        335..613
FT                  /label= 29kD protein incl. AR DNA-binding domain
FT   Region         614..630
FT                  /label= 17 amino acid linker
XX
PN   WO9107423-A.
XX
PD   30-MAY-1991
XX
PF   19-OCT-1990;  90WO-US06015.
XX
PR   17-NOV-1989;  89US-0438775.
XX
PA   (ARCH-) ARCH DEV CORP.
```

```
XX
PI   Liao S, Chang C;
XX
DR   WPI; 1991-178048/24.
DR   N-PSDB; AAQ12008.
XX
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
XX
PS   Example 13; Fig 10; 79pp; English.
XX
CC   To express an androgen receptor fusion protein in E.coli, the pATH
CC   expression system was used. The trpE is insoluble so partially
CC   purified induced fusion protein is obtained by simply lysing the
CC   E.coli and precipitating the insoluble fusion protein. The fusion
CC   protein was used for immunisation to obtain monoclonal anti-AR
CC   antibodies.
XX
SQ   Sequence    630 AA;

Query Match          29.7%;   Score 1459; DB 12;  Length 630;
Best Local Similarity 79.9%;   Pred. No. 3.1e-90;
Matches   286;   Conservative  17;  Mismatches  31;  Indels  24;  Gaps  6;

Qy   515  VPYPSPTCV----KSEMGPWM----DS----YSGPYGDMRLE-TARDHVLPIDYFPP 559
          :|  ||   :|  :|:|      |:|   :: :|  :| :|:|
Db   272  lpcpsplaayyvlkksnpspymfimqdndftlfgaspesslkydatsrqieipgrarlra 331

Qy   560  QKTCLICGDEASGCHYGALTCGSCKVFFKRAAEGKQKYLCASRNDCTIDKFRRKNCPSCR 619
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   332  rpg--icgdeasgchygaltcgsckvffkraaegkqkylcasrndctidkfrrkncpscr 389

Qy   620  LRKCYEAGMTLGARKLKKLGNLKLQEEGEASSTTSPTEETTQKLTVSHIEGYECQPIFLN 679
          ||||||||||||||||||||||| || :||:||||||||||  ||||||||||||||||
Db   390  lrkcyeagmtlgarklkklgnlklqeegenssagspteedpsqkmtvshieqyecqpifln 449

Qy   680  VLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMA 739
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   450  vleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwakalpgfrnlhvddqma 509

Qy   740  VIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFG 799
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   510  viqysxmglmvfamgwrsftnvnsrmlyfapdlvfneyrmhksrmysqcvrmrhlsqefg 569

Qy   800  WLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCS 857
          ||||||||||||||||||||||||||||||||||||||||||   :||
Db   570  wlqitpqeflcmkavllfsiipvdglknqkffdelrmnyikeldpl------estcs 620
```

```
RESULT 12
AAB97073
ID   AAB97073 standard; Protein; 260 AA.
XX
AC   AAB97073;
XX
DT   31-JUL-2001  (first entry)
XX
DE   Rat androgen receptor ligand binding domain.
XX
KW   Rat; androgen receptor; AR; ligand binding domain; LBD; osteopathic;
KW   crystallographic structure; AR-LBD; AR modulator; prostate cancer;
KW   age related disease; osteoporosis; muscle wasting; libido; vasotropic;
KW   protein coordinate data.
XX
OS   Rattus sp.
XX
PN   WO200127622-A1.
XX
PD   19-APR-2001.
XX
PF   13-OCT-2000; 2000WO-US28495.
XX
```

```
PR   14-OCT-1999;    99US-0159394.
XX
PA   (BRIM)  BRISTOL-MYERS SQUIBB CO.
XX
PI   Weinmann R, Einspahr HM, Krystek SR, Sack JS, Salvati ME;
PI   Tokarski JS, Wang C, Attar RM;
XX
DR   WPI; 2001-300222/31.
XX
PT   New crystallographic structure of the Androgen receptor ligand binding
PT   domain, useful for identifying modulators of androgen receptors -
XX
PS   Example; Page 27; 83pp; English.
XX
CC   The present sequence is the rat androgen receptor (AR) ligand-binding
CC   domain (LBD). It is provided in an example illustrating an invention
CC   relating to a new crystallographic structure of the AR-LBD. The
CC   structure comprises either an AR-LBD and an AR-LBD ligand, or an
CC   AR-LBD without an AR-LBD ligand (where the crystal diffracts to at
CC   least 3 angstrom resolution and has a crystal stability within 5 % of
CC   its unit cell dimensions). The AR-LBD is useful for identifying
CC   modulators of the androgen receptor, which may be used for treating
CC   prostate cancer and age related diseases such as osteoporosis, muscle
CC   wasting and loss of libido.
XX
SQ   Sequence    260 AA;

Query Match          27.5%;   Score 1349.5; DB 22;  Length 260;
Best Local Similarity 99.6%;   Pred. No. 2.5e-83;
Matches   258;   Conservative   0;  Mismatches   0;  Indels   1;  Gaps   1;

Qy   666  SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 724
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2    shmiegyecqpiflnvleaiepgvvcaghdnnqpdsfaallsslnelgerqlvhvvkwak 61

Qy   725  ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62   alpgfrnlhvddqmaviqyswmglmvfamgwrsftnvnsrmlyfapdlvfneyrmhksrm 121

Qy   785  YSQCVRMRHLSQEFGWLQITPQEFICMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122  ysqcvrmrhlsqefgwlqitpqeficmkalllfsiipvdglknqkffdelrmnyikeldr 181

Qy   845  IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182  iiackrknptscsrrfyqltklldsvqpiarelhqftfdllikshmvsvdfpemmaeiis 241

Qy   905  VQVPKILSGKVKPIYFHTQ 923
          ||||||||||||||||||
Db   242  vqvpkilsgkvkpiyfhtq 260
```

```
RESULT 13
AARI2229
ID   AARI2229 standard; Protein; 576 AA.
XX
AC   AARI2229;
XX
DT   20-AUG-1991  (first entry)
XX
DE   TrpE/androgen receptor N-terminal domain fusion protein.
XX
KW   androgen receptor; AR; DNA-binding protein; steroid hormone.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..323
FT                   /label= 33kD trpE protein
FT   Region          324..329
FT                   /label= 6 amino acid linker
FT   Protein         330..571
XX
```

```
FT   Region          /label= N-terminal region of AR
FT                   572..576
FT                   /label= 5 amino acid linker
XX
PN   WO9107423-A.
PD   30-MAY-1991.
XX
PF   19-OCT-1990;  90WO-US06015.
XX
PR   17-NOV-1989;  89US-0438775.
XX
PA   (ARCH-) ARCH DEV CORP.
XX
PI   Liao S,  Chang C;
XX
DR   WPI; 1991-178048/24.
DR   N-PSDB: AAQ12007.
XX
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
XX
PS   Example 13; Fig 9; 79pp; English.
XX
CC   To express an androgen receptor fusion protein in E.coli, the pATH
CC   expression system was used. The trpE is insoluble so partially
CC   purified induced fusion protein is obtained by simply lysing the
CC   E.coli and precipitating the insoluble fusion protein. The fusion
CC   protein was used for immunisation to obtain monoclonal anti-AR
CC   antibodies.
XX
SQ   Sequence  576 AA;

Query Match            27.2%;  Score 1337.5;  DB 12;  Length 576;
Best Local Similarity  55.5%;  Pred. No. 4.4e-82;
Matches 294;  Conservative  39;  Mismatches  92;  Indels 105;  Gaps  17;

Qy   106 PTGYLVLDEEQQPS-------QPQSALECHPERGCVPEPGAAVAASKGLPQQLPAPPDEDD 159
            |   ||   ||   |           | | ||     | |        |  ||| : ||:
Db    89 pag---veseqspncrvlrfppvsplldedarlc----slsvfdafrllqnllnvpkeer 141

Qy   160 SAAPSTLSLLGPTFPGLSSCS-----ADLKDILSE--------ASTMQLLQQQQQEAV 204
            | |    || |   | | |         ||  ||             |: |: ||
Db   142 eam-------ffsglfsydlvagfedlpqlsaenncpdfcfylaetlmvidhqkk--- 189

Qy   205 SEGSSSGRAREASGAPISSKDNYLGGTSTISDNAKELCKAVSVSMGLGVEALEHLSPGEQ 264
            |   | |||| : :     :  |: |: | :   :  :  : | :|  ||  |:
Db   190 ----striqaslfapneeekqrl--tarlnelrqqltea---applpvvsvph------- 233

Qy   265 LRGDCMYA--------------PLLGVPPAVR------PTPCAPLAECKGS----- 295
            :| : |              ||   | | |        | | ||| ||
Db   234 mrcecnqsdeefggvvrllqkalrageifqvvpsrrfslpcpsplaayyvlkksnpspym 293

Qy   296 -LLDDS----AGKSTEDTAEYSPFKGGYTKGLEGESLGCSSAAAGSSGTLELPSTLSLY 350
             ||::          | :  |   |      |          :   ||| |  | |:|
Db   294 ffmqdndftlfgaspesslkyd---atsrqie--------iprnsssgtlelpstlsly 341

Qy   351 KSGALDEAAAYQSRDYNFPLALAGPPPPPPHPHARIKLENPLDYGSAWAAAAAQCRY 410
            |||||| | | |||||| | |||||||||||||||||||||||||||  | |||||
Db   342 ksgaldeaaayqsrdyynfplalagpppppppphphariklenpldygsawaaaaaqcry 401

Qy   411 GDLASLHGAGAAGPGSGSPSAAAASSWHTLFTAEEGQLYGPC--GGGGGGGGGGGGG 467
            ||||||||| ||||||||||||||||||||||:  || | ||  ||||||||||||||
Db   402 gdlaslhgagaagpgsgspsaaaasswhtlftaeegqlyqpcgggggggggggggggg 461

Qy   468 GGGGGGGGGGEGAVAPYGYTRPPQGLAGQESDFTAPDVWTPGGMVSRVPYPSPTCVKSEM 527
            ||||||||| | |  || ||||||  || : :  |  | || |  |:| |||||| ||||
Db   462 ggggggggggeaeavapygytrppqqglaggqesdftapdvwypggmvsrvpypsptcvksem 521

Qy   528 GPWMDSYSGPYGDMRLETARDHVLPIDYFPPPQKTCLICGDEASGCHYGA 577
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   522 gpwmdsysgpyqdmrletardhvlpidyyfppqktclicgdeasgchyga 571
```

```
RESULT 14
AAY21621
ID   AAY21621 standard; protein; 933 AA.
XX
AC   AAY21621;
XX
DT   11-AUG-1999  (first entry)
XX
DE   Ligand binding domain of nuclear receptor hPR.
XX
KW   Thyroid hormone receptor; aromatic compound; ligand binding domain;
KW   alpha-glycerophosphate dehydrogenase; cardiac; obesity; triglyceride;
KW   plasma cholesterol; anti-hypertriglyceridaemic; atherosclerosis; GPDH;
KW   thyroid hormone replacement therapy; nuclear receptor.
XX
OS   Homo sapiens.
XX
PN   WO9926966-A2.
XX
PD   03-JUN-1999.
XX
PF   25-NOV-1998;  98WO-US25296.
XX
PR   26-NOV-1997;  97US-0980115.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Apriletti JW,  Baxter JD,  Fletterick RJ,  Kushner PJ;
PI   Scanlan TS,  Shiau AK,  Wagner RL,  West BL;
XX
DR   WPI; 1999-357810/30.
XX
PT   Modulating activity of a thyroid hormone receptor
XX
PS   Disclosure; Fig 3A-R; 447pp; English.
XX
CC   The invention relates to a method for modulating activity of a thyroid
CC   hormone receptor that comprises administration of an aromatic compound
CC   which fits spatially and preferentially into a thyroid hormone ligand
CC   binding domain. The aromatic compound (of a specified formula) can be
CC   used to increase alpha-glycerophosphate dehydrogenase (GPDH) levels, at
CC   levels which do not significantly modify cardiac GPDH levels and are
CC   indicated in the treatment of obesity. The compound also lower total
CC   plasma cholesterol and triglyceride levels and can be used as anti-
CC   hypertriglyceridaemic agents. The compound may also be used for treating
CC   atherosclerosis and may be indicated in thyroid hormone replacement
CC   therapy in patients with compromised cardiac function. Sequences
CC   AAY21621- 636 amino acid sequences of ligand binding domains of several
CC   members of the nuclear receptor superfamily.
XX
SQ   Sequence  933 AA;

Query Match            25.7%;  Score 1261;  DB 20;  Length 933;
Best Local Similarity  34.1%;  Pred. No. 1.2e-76;
Matches 346;  Conservative 128;  Mismatches 318;  Indels 224;  Gaps  34;

Qy    38 GPRHPEAASAAP----------------------PGA----------------SLLLQQQQQQQ 66
            |||  ||  || |                     ||                  |||||||||||
Db     8 gpraphvaggppspevgspllcrpaagpfpgsqtsdtlpevsaipisldgllfrprpcqgg 67

Qy    67 QQQQQQQQQQQETSPRQQQQQGEDGSPQAHRRGPTGYL--VLDEEQQPSQPQSA 124
            : :: : |          :  ::  : : | :  | ::: ||||||||||||||||
Db    68 dpsdektqdqqslsdvegaysraeatrgaggsssspppekdsglldsvldtllapsgpggs 127

Qy   125 LECHPERGCVPEPGAAVAASKG--LPQQLPAPPDEDDSAAPSTLSLLGPTFP------GL 176
            |  :  | |  : |::: : ||   |:  : |  : |   |  |   | | |       ||
Db   128 --------qpsppacevtsswclifgpelp---edppaapatqrvlsplmsrsgckvgd 174

Qy   177 SSCSADLKDILSE--ASTMQLL------ 218
            |  ||||  | |   |||  ||
Db   177 QQQQQEAVS-EGSSSGRAREASG 218
```

PN    WO200049147-A1
XX
PD    24-AUG-2000.
XX
PF    18-FEB-2000; 2000WO-EP01368.
XX
PR    19-FEB-1999;  99DE-1007099.
PR    19-FEB-1999;  99US-0120848.
XX
PA    (THER-) THERAGENE BIOMEDICAL LAB GMBH.
XX
PI    Hauser-funke C;
XX
DR    WPI; 2000-549273/50.
DR    N-PSDB; AAA53851.
XX
PT    Novel nucleic acid construct useful in gene therapy comprising an
PT    hormone responsive element and transgene in which the hormone
PT    responsive element is not functionally linked to the transgene
XX
PS    Disclosure; Page 92-95; 100pp; English.
XX
CC    New nucleic acid constructs are described which comprise an hormone
CC    responsive element (HRE) and a transgene (T). Alternatively the
CC    nucleic acid construct, comprises at least one HRE and a transgene,
CC    where one of the HREs is not functionally linked to the transgene;
CC    The constructs can be used to up-regulate or down-regulate target
CC    genes and for the delivery of vaccines. The constructs preferably
CC    comprise a transgene which encodes a protein which is lacking in a
CC    variety of genetic disorders or involved in conditions related to
CC    inappropriate responses to hormones, for example hormone-dependent
CC    cancers such as breast, ovarian, and endometrial cancers and prostate
CC    cancer. The transgene may also be used to replace a defective gene
CC    resulting in such genetic disorders as haemophilia, von Willebrand
CC    disease, and cystic fibrosis. Vectors comprising these constructs
CC    where the transgene is human clotting factor IX can be used for
CC    treating blood clotting disorders such as haemophilia A or B on
CC    administration to an organism or to a cellular system. The constructs
CC    have applications in gene therapy for treating haemophilia when the
CC    transgene encodes a clotting factor such as clotting factor IX. The
CC    advantage of this systme is that the hormone-hormone receptor complex
CC    contains a hormone receptor that becomes activated after binding of
CC    its specific hormone. The hormone receptor in the activated state is
CC    able to recognise and bind to its specific hormone responsive
CC    element. The presence of the hormone responsive element on the
CC    nucleic acid carrying a transgene encourages binding of a
CC    hormone-hormone receptor complex. Thus the activated hormone receptor
CC    acts as a link between the nucleic acid carrying the transgene and the
CC    hormone known to interact with the cell membrane.
XX
SQ    Sequence  933 AA;

Query Match                25.6%;   Score 1256;   DB 21;   Length 933;
Best Local Similarity       34.5%;   Pred. No. 2.5e-76;
Matches 349;   Conservative 119;   Mismatches 329;   Indels 214;   Gaps 33;

Qy    38   GPRHPEAASAAP---------------------PGA--------------SLLLQQQQQQQQ    66
                    ||||    |        ||:                              ||  |
Db     8   gpraphvaggpppspevgspllcrpaagpfpgsqtsdtlpevsalpisldgllfprpcggq    67

Qy    67   QQQQQQQQQQQQQQQQQQQETSPROQQQQQGEDGSPQAHRRGPTGYL--VLDEEQQPSQPQSA  124
                                    :||    ||||:    :|| ::|  |::|| | |: :|:
Db    68   dpsdektqdqgslsdvegaysraeatrgaggsssppekdsglldsvidtllapsgpgqs    127

Qy   125   LECHPERGCVPEPGAAVAASKG--LPQQLPAPPDEDDSAAPSTLSLLGPTFP------GL  176
              ||       |||||||||||     ||    |    ||  ||  || :|:|||||
Db   128   ------qpsppacevtsswclfgpelp---edppaapatqrvlplmsrsgckvyd  174

Qy   177   SSCSADLKDILSE--ASTMOLL-----QQQQQEAVSEGSSSGRAREASGA  219
              |  |:|   ||:|    :|||
Db   175   ssgtaaahkvlprglsparqlllpasesphwsgapvkpspqaaaveveedgseseesag  234

Db   175   ssgtaaahkvlprglsparqlllpasesphwsgapvkpspqaaaveveedgseseesag  234
Qy   219   APTSSKDNYLGGTSTISDNAKELCKAVSVSMGLG------------------VEALEHLSPG  262
                   :||   || |     |                                 ::||
Db   235   pllkgkpralggaa--agggaaacppgaaaggvalvpkedsrfsaprvalveqdapmapg  292
Qy   263   EQLRG----DCMYAPLLGV----------------------------PPAVRPTPCA--  287
                 |:|      ::|:|                                ||   :|||
Db   293   rsplattvmdfihvpilplnhallaartrqlledesydggagaasafapp--rtspccass  350
Qy   288   -PLAECKGSLLDDSAGKSTEDTAE----YSPFKGGYTKGLEGESLGCSGSAAAGSSGTLE  342
             :|      |     |  ||| | |:   :|  ||::| : :|  |:  ||||
Db   351   tpva--vgdfpdcayppdaepkddayplysdfqppalk-ikeeeegaeasarsprs----  403
Qy   343   LPSTLSLYKSGALDEAAAYQSRDYNFPLALAGPPPPPPPHPHARIKLENPLDYGSAWA  402
               :    ||                  |||       || |  | ||
Db   404   -------ylvaganpaa-------fpdfpl--gpppplpp------------------  427
Qy   403   AAAAQCRYGDLASLHGAGAGPGSGSPSAAAASSSWH--TLFTAE---EGQLYGPCGG  454
              |                   || |||      |:| |        ||  |:  |
Db   428   -ratpsrpge---aavtaapasavssassgstleclikaegappqqgpfapppck  481
Qy   455   GGGGGGGGGGGGGGGGGGGGEAGAVAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSR  514
              |||||||||||||||||||| |   ||| |  |  |  |         ::
Db   482   apgasgcllprdglpstsasaaaqa-apaly-palging------lpqlgyyqaavike  532
Qy   515   -VPYFSPTCVKSEMGPWMDSYSGPYGDMRLETARDHVLPIDYYFP--PQKTCLICGDEAS  571
              |  ||     |  |:  :|| |   |   ||  | | | :||     ||:|  |||:|
Db   533   glpqvyppyl-nylrpdseasqsp-----------qysfeslpqkiclicgdeas  575
Qy   572   GCHYGALTCGSCKVFFKRAAEGKQKYLCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLG  631
              |||||||||||||||||||||| |:  |||| |||  |||||||||| |:: |||| ||
Db   576   gchygvltcgsckvffkramegqhylcagrndcivdkirrkncpacrlrkccqagmvlg  635
Qy   632   ARKLKLGNLKLQERGEASSTTSP------TEETTQKLTVSHIEGYECQPIFLNVLEAIE  685
              || |  |  ||:||| |: |  |        : |  | |  || :  |:|  | |:|:|
Db   636   grkfkfknkvvvraldaval-pqplgvpnesqalsqrftfspgqdiqllpplinllmsle  695
Qy   686   PGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMAVIQYSW  745
              |  | |||  | |||  || |||| ||||| | || ||| ||| ||||| |  : |:|||
Db   696   pdviyaghdntkpdtssllstslnqlgerqlisvvkwsksplpgtrnlhiddqitliqysw  755
Qy   746   MGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQITP  805
              :| |||:| |  |:  ||  | : |||:|  |: | |:|:  |  | ::|||| |: ||
Db   756   mslmvfiglwrtsykhvsgqmlyfapdlllneqrmkessfyslcltmwqipqefvklqvsq  815
Qy   806   QEFLCMKALLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRRFYQLTK  865
              :|| | |    :|     |::||| | | ||| |  |||  |  |:| | ||||||||
Db   816   eeflcmkvlllintipleglrsqtqfeemrssyirelikaiglrqkgvvsssqrfyqltk  875
Qy   866   LLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIYFH  921
              ||| :|| |  |: |  :||:  |     || :|::||: | |  |  |||||||
Db   876   lldnlhdlvkglhlyclntfiqsralsvefpemnseviaaglpkilagmvkpllfh  931

RESULT  15
AAY97297
ID    AAY97297 standard; Protein; 933 AA.
XX
AC    AAY97297;
XX
DT    03-JAN-2001 (first entry)
XX
DE    Human progesterone receptor B-form.
XX
KW    Recombinant DNA; gene therapy; hormone responsive element;
KW    transgene; HRE; haemophilia; clotting factor IX; vaccine;
KW    regulation; breast cancer; ovarian cancer; prostate cancer;
KW    von Willebrand disease; cystic fibrosis; hormone; receptor; human;
KW    blood.
XX
OS    Homo sapiens.
XX

us-09-497-822a-19.rag

```
Qy   220 P-TSSKDNYLGGTSISDNA------------------KELCKAVSVSMGLGVEALEHLSPGE 263
             |   :| |||  ||            ||:  : :|| :  : :||
Db   235 pllkqkpralggaaaggpaaaavppgaaaggvalvpkedsrfsaprval-veqdapmapgr 293

Qy   264 QLRG---DCMYAPLLGVPPAVRPTPCAPLAECKGSLLDDSAGKSTEDTAEYSPFKGGYT 319
             :: |  :|:|           || ||                            :
Db   294 splattvmdfihvpil------plnhallaartrqlledes--------ydggaqaas 337

Qy   320 KGLGEGESLGCSGS--AAAGSSGTLELP------STLSLY-----------KSGALDE 357
             |:: |  |  |            ||              ||
Db   338 afapprsspcasstpvavgdfpdcayppdaepkddayplysdfqppalkikeeeqa--e 395

Qy   358 AAAYQSRDY---------NFPLALAGPPPPPPHPHARIKLENPLDYGSAWAAAAAQ 407
             |:| | |         |||| |||| ||
Db   396 asarsprsylvaganpaafpdfpl--gpppplpp------------------ratp 431

Qy   408 CRYGDLASLHGAGAAGPGSGSEPSAAASSSWH---TLFTAE-----EGQLYGPCGGGGGGGG 459
             |||  || |  |: |||||||||| |||   :|          :|
Db   432 srpge----aavtaapasasvssasssgstlecilykaegappqqgpfappppckapgas 486

Qy   460 GGGGGGGGGGGGGGGGGEAGAVAPYGYTRPPQQGLAGQESDFTAPDVWYPGGMVSR-VPYP 518
             || ||                           |       | : :: |
Db   487 gcllprdglpstsasaaaga-apaly--palging-----lpqigyqaavlkeglpqv 537

Qy   519 SPTCVKSEMGPWMDSYSGPYGDMRLETARDHVLPIDYYFP--PQKTCLICGDEASGCHYG 576
             | :: |                   |  |:| |||:||||||||||
Db   538 yppyl-nylrpdseasqsp-----------qysfeslpqkiclicgdeasgchyg 580

Qy   577 ALTCGSCKVFFKRAAEGKQKYLCASRNDCTIDKFRRKNCPSCRLRKCYEAGMTLGARKLK 636
             ||||||||||||||| ||:  :||: ||||:|| :|||:  ::||: :||| || |||||
Db   581 vltcgsckvffkramegqhnylcagrndcivdkirrkncpacrlrkccqagnvlggrkfk 640

Qy   637 KLGNLKLQEEGEASSTTSP-----TEETTQKLTVSHIEGYECQPIFLNVLEAIEPGVVC 690
             :: :|  |       ::  : |:  | :  :|  :|| ::|||  | :| :||
Db   641 kfnkvrvvraldavalpqplgvpnesqalsqrftfspgqdiqlippliniilmsiepdviy 700

Qy   691 AGHDNNQPPDSFAALLSSLNELGERQLVHVKWAKALPGFRNLHVDDQMAVIQYSWMGLMV 750
             ||||||| :::  :||:||||||| ||: ||||| |::|||||::|::: :||||| |||
Db   701 aghdntkpdtssslltslnqlgerqlisvkwkslpqfrnlhiddqitliqyswmslmv 760

Qy   751 FAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLC 810
             | :: ||| || :|||||||||: :||| |:: | | |:  || |:: :|::: || |
Db   761 fglgwrsykhvsgqmlyfapdlilneqrmkessfysicltmwqipqefvklqvsqeflc 820

Qy   811 MKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRRFYQLTKLLDSV 870
             |: :||  | :::: | :||:: : :| ||: :| : | | || |:|||:||:|||||:::
Db   821 mkvlllntipleglrsqtqfeemrssyireliikaiglrqkgyvsssqrfyqltklldni 880

Qy   871 QPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIYFH 921
             |:|  :||:|  :||::||||:||| |:| ||:| | | |||||||
Db   881 hdlvkqlhlyclntfiqsralsvefpemmseviaaqlpkilagmvkpllfh 931
```